EXPONENTIAL DOWNWEIGHTING OF PAST DATA IN A SINGLE-STAGE, WEIGHTED-LEAST-SQUARES TRAJECTORY PROCESSOR

by
Robert G. Rich
Trajectory Determination Department

Approved by

Alfred H. Kircher, Jr.
Manager, Trajectory Determination Department

Submitted to

National Aeronautics and Space Administration
Manned Spacecraft Center
Houston, Texas 77058

Contract No. NAS 9-996

Federal Systems Division
International Business Machines Corporation
1322 Space Park Drive
Houston, Texas 77058

70-5667
12/16/70

EXPONENTIAL DOWNWEIGHTING OF PAST DATA IN A SINGLE-STAGE, WEIGHTED-LEAST-SQUARES TRAJECTORY PROCESSOR

INTRODUCTION

Trajectories of coasting spacecraft usually are estimated by processing radar data through some sort of weighted-least-squares filter. Where trajectory model errors are of practical significance, the measurements are processed sequentially in small batches as received, and the state covariance matrix is adjusted during propagation between stages. This procedure is successful as long as the navigational computer is continually available. For the future, however, long-duration missions are being planned with a new requirement for storing the measurements and processing them much less frequently. For example, a single-stage filter might be used at 24-hour intervals to process, in each case, the past 15 hours of stored data. The accuracy of each estimate, of course, would depend on the magnitude of model errors along the data arc. This is the motivation for seeking a way of adjusting the single-stage, weighted-least-squares filter to compensate for model errors. The following empirical method does this by allowing exponential downweighting of past data at selected rates along respective axes of a selected coordinate system.

Notation and Definitions:

1. IC — inertial coordinate(s)

2. RC — rotating coordinate(s)

3. Superscripts:
   - $\hat{}$ — unit vector
   - Δ — better approximation
   - $\bar{}$ — position or velocity vector
   - $\tilde{}$ — a variable held constant; also the measured value of an observation
   - T — transpose of matrix or vector

4. $\emptyset$ — null matrix or vector

5. $\bar{r}^T \begin{cases} \equiv [x \;\; y \;\; z\,] \text{ (IC)} \\ \equiv [u_1 \;\; u_2 \;\; u_3] \text{ (RC)} \end{cases}$ position vector (transpose)

6. $\bar{v}^T \begin{cases} \equiv [\dot{x} \;\; \dot{y} \;\; \dot{z}\,] \text{ (IC)} \\ \equiv [u_4 \;\; u_5 \;\; u_6] \text{ (RC)} \end{cases}$ velocity vector (transpose)

7. $\hat{r}$ — unit position vector

8. $\hat{v}$ — unit velocity vector

9. $\hat{w} \equiv \dfrac{\bar{r} \times \bar{v}}{|\bar{r} \times \bar{v}|}$

10. $\hat{u} \equiv \hat{w} \times \hat{r}$

11. $\hat{q} \equiv \hat{v} \times \hat{w}$

The orthonormal basis vectors for the selected RC are $\{\hat{e}_j | j=1,\ldots,6\}$:

12. $\hat{e}_1 \equiv \begin{bmatrix} \hat{r} \\ \emptyset \end{bmatrix}$; $\hat{e}_2 \equiv \begin{bmatrix} \hat{u} \\ \emptyset \end{bmatrix}$; $\hat{e}_3 \equiv \begin{bmatrix} \hat{w} \\ \emptyset \end{bmatrix}$ (6×1)

13. $\hat{e}_4 \equiv \begin{bmatrix} \emptyset \\ \hat{q} \end{bmatrix}$; $\hat{e}_5 \equiv \begin{bmatrix} \emptyset \\ \hat{v} \end{bmatrix}$; $\hat{e}_6 \equiv \begin{bmatrix} \emptyset \\ \hat{w} \end{bmatrix}$ (6×1)

(RC at time t_0 is the selected system for downweighting.)

14. t_0 anchor time (time of last observation in the batch)

15. t_i time of the ith observation $(i=1,\ldots,n)$ in a batch of n observations where $t_n = t_0$

16. $\begin{bmatrix} \bar{r}_0 \\ \bar{v}_0 \end{bmatrix}$ state vector at time t_0

17. $S_0^T \equiv [x\ y\ z\ \dot{x}\ \dot{y}\ \dot{z}]_0$ state vector (transpose) at time t_0, expressed in IC

18. $U_0^T \equiv [u_1 \ldots u_6]_0$ state vector (transpose) at time t_0, expressed in RC

19. $E_0 \equiv [\hat{e}_1 \ldots \hat{e}_6]_0$ the matrix of RC basis vectors at time t_0 (12,13)

(Note that $E_0^{-1} = E_0^T$.)

Then

20. $S_0 = [\hat{e}_1 u_1 + \cdots + \hat{e}_6 u_6]_0 = E_0 U_0$ (18,19)

21. $\alpha_j \geq 1$ $(j=1,\ldots,6)$ chosen scalar constants (the downweighting parameters)

22. $A_{0i} \equiv \begin{bmatrix} \alpha_1^{t_i - t_0} & & \emptyset \\ & \ddots & \\ \emptyset & & \alpha_6^{t_i - t_0} \end{bmatrix}$ (21)

23. $\Omega_{0i} \equiv A_{0i} E_0^T$ (19,22)

24. $\tilde{\beta}_i$ the observed value of a measurement vector at time t_i

25. $\beta_i = \beta_i(u_1,\ldots,u_6)_0$ the measurement vector at time t_i, computed as a function of the independent variables in RC at time t_0

26. $\delta\beta_i \equiv \tilde{\beta}_i - \beta_i$ vector of residuals

27. $\tilde{u}_j$ an RC independent variable held constant (3,18)

28. $\tilde{U}_0^T = [\tilde{u}_1\ldots\tilde{u}_6]_0$ an initial approximation of the state vector at time t_0, expressed in RC (18,27)

Derivation:

Let every β_i be associated with $\{\beta_{ij}|j=1,\ldots,6\}$ as

29. $$\beta_i = \beta_i(u_1,\ldots,u_6)_0 \longleftrightarrow \left\{\begin{matrix} \beta_{i1} \equiv \beta_i(u_1,\tilde{u}_2,\ldots,\tilde{u}_6)_0 \\ \vdots \\ \beta_{i6} \equiv \beta_i(\tilde{u}_1,\ldots,\tilde{u}_5,u_6)_0 \end{matrix}\right\} .$$

Also, similar to 26,

30. $$\delta\beta_{ij} \equiv \tilde{\beta}_i - \beta_{ij} .$$

Partial derivatives of β_i and β_{ij} with respect to U_0 are

31. $$\frac{\partial\beta_i^T}{\partial U_0} = \begin{bmatrix} \dfrac{\partial\beta_i^T}{\partial u_1} \\ \vdots \\ \dfrac{\partial\beta_i^T}{\partial u_6} \end{bmatrix}$$ (18,25)

32. $$\sum_{j=1}^{6} \frac{\partial\beta_{ij}{}^T}{\partial U_0} = \begin{bmatrix} \dfrac{\partial\beta_{i1}{}^T}{\partial u_1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \dfrac{\partial\beta_{i6}{}^T}{\partial u_6} \end{bmatrix} \quad (6 \times 1)$$ (18,28,29)

Then

33. $$\frac{\partial \beta_i{}^T}{\partial U_0} = \sum_{j=1}^{6} \frac{\partial \beta_{ij}{}^T}{\partial U_0} \; . \qquad (31,32)$$

Also define

34. W_i , the weighting matrix associated with both $\delta\beta_i$ and $\delta\beta_{ij}$.

The quadratic form to be minimized in the weighted-least-squares process is

35. $$2\varphi \equiv \sum_{i=1}^{n} \sum_{j=1}^{6} \delta\beta_{ij}{}^T \; \alpha_j{}^{t_i - t_0} \; W_i \; \delta\beta_{ij} \; .$$

36. $$\Phi \equiv \frac{\partial \varphi^T}{\partial U_0} = - \sum_{i=1}^{n} \sum_{j=1}^{6} \frac{\partial \beta_{ij}{}^T}{\partial U_0} \; \alpha_j{}^{t_i - t_0} \; W_i \; \delta\beta_{ij} = \emptyset \qquad (35)$$

Assume that equation 36 has a unique solution. Since 35 is positive definite, it will be minimized by the solution to 36. Express the solution to 36 in RC as $\overset{\Delta}{U}_0$. This solution $\overset{\Delta}{U}_0$ depends on the initial approximation $\tilde{U}_0$, which remains constant (27,28). However, we want the particular solution such that

37. $$\overset{\Delta}{U}_0 = \tilde{U}_0 \; .$$

Condition 37 implies that

38. $$\delta\beta_{ij} = \delta\beta_i \qquad (j=1,\ldots,6) \; . \qquad (29)$$

Now the summation on j can be removed, and equation 36 can be simplified as follows:

$$\Phi = - \sum_{i=1}^{n} \sum_{j=1}^{6} \frac{\partial \beta_{ij}{}^T}{\partial U_0} \; \alpha_j{}^{t_i - t_0} \; W_i \; \delta\beta_i \qquad (36,38)$$

$$= - \sum_{i=1}^{n} \sum_{j=1}^{6} A_{0i} \frac{\partial \beta_{ij}{}^T}{\partial U_0} W_i \quad \delta\beta_i \qquad (22,32)$$

$$= - \sum_{i=1}^{n} A_{0i} \frac{\partial \beta_i{}^T}{\partial U_0} W_i \quad \delta\beta_i \qquad (33)$$

$$= - \sum_{i=1}^{n} A_{0i} \frac{\partial S_0{}^T}{\partial U_0} \frac{\partial \beta_i{}^T}{\partial S_0} W_i \quad \delta\beta_i$$

$$= - \sum_{i=1}^{n} A_{0i} E_0^T \frac{\partial \beta_i{}^T}{\partial S_0} W_i \quad \delta\beta_i \qquad (20)$$

39. $$\Phi = - \sum_{i=1}^{n} \Omega_{0i} \frac{\partial \beta_i{}^T}{\partial S_0} W_i \quad \delta\beta_i = \emptyset \qquad (23)$$

The solution to equation 39 will minimize 35 under condition 37. Computationally it is convenient to express the solution to equation 39 in IC, i.e., find $\overset{\Delta}{S}_0$ rather than $\overset{\Delta}{U}_0$. This is done by successive, nearly-linear approximations (neglecting second order partials) as follows:

In equation 39 assume $\Phi(S_0) \neq \emptyset$, where S_0 is an initial estimate of the solution. Let $\overset{\Delta}{\Phi} = \Phi(\overset{\Delta}{S}_0) = \emptyset$. Then a linear approximation of $\overset{\Delta}{\Phi}$ is

40. $$\overset{\Delta}{\Phi} \approx \Phi + \frac{\partial \Phi}{\partial S_0} (\overset{\Delta}{S}_0 - S_0) = \emptyset \quad ,$$

from which

41. $$\overset{\Delta}{S}_0 = S_0 - \left(\frac{\partial \Phi}{\partial S_0}\right)^{-1} \Phi \quad .$$

The approximate value of $\frac{\partial \Phi}{\partial S_0}$ (neglecting second-order partials) is

42. $$\frac{\partial \Phi}{\partial S_0} \approx \sum_{i=1}^{n} \Omega_{0i} \frac{\partial \beta_i{}^T}{\partial S_0} W_i \frac{\partial \beta_i}{\partial S_0} , \qquad (39)$$

and letting $\delta \overset{\Delta}{S}_0 \approx \overset{\Delta}{S}_0 - S_0$,

43. $$\delta \overset{\Delta}{S}_0 = \left[\sum_{i=1}^{n} \Omega_{0i} \frac{\partial \beta_i{}^T}{\partial S_0} W_i \frac{\partial \beta_i}{\partial S_0} \right]^{-1} \left[\sum_{i=1}^{n} \Omega_{0i} \frac{\partial \beta_i{}^T}{\partial S_0} W_i \, \delta \beta_i \right] .$$

We assume in equation 43 that S_0 is in the region of convergence and that the matrix inverse exists. Then the solution $\overset{\Delta}{S}_0$ is obtained by iterating until convergence criteria are satisfied:

44. $$0 \le \delta \overset{\Delta}{S}{}_0^T \, \delta \overset{\Delta}{S}_0 < \varepsilon \text{ chosen} \qquad (43)$$

Equation 43, then, is the modified filter.

If $\alpha_j = \alpha$ $(j=1,\ldots,6)$ (22), then equation 43 simplifies to

45 $$\delta \overset{\Delta}{S}_0 = \left[\sum_{i=1}^{n} E_0^T \frac{\partial \beta_i{}^T}{\partial S_0} \alpha^{t_i - t_0} W_i \frac{\partial \beta_i}{\partial S_0} \right]^{-1} \left[\sum_{i=1}^{n} E_0^T \frac{\partial \beta_i{}^T}{\partial S_0} \alpha^{t_i - t_0} W_i \, \delta \beta_i \right]$$

or equivalently, since $E_0^{-1} = E_0^T$,

46. $$\delta \overset{\Delta}{S}_0 = \left[\sum_{i=1}^{n} \frac{\partial \beta_i{}^T}{\partial S_0} \alpha^{t_i - t_0} W_i \frac{\partial \beta_i}{\partial S_0} \right]^{-1} \left[\sum_{i=1}^{n} \frac{\partial \beta_i{}^T}{\partial S_0} \alpha^{t_i - t_0} W_i \, \delta \beta_i \right]$$

If further, $\alpha_j = 1$ $(j=1,\ldots,6)$, then

47. $$\delta \overset{\Delta}{S}_0 = \left[\sum_{i=1}^{n} \frac{\partial \beta_i^T}{\partial S_0} W_i \frac{\partial \beta_i}{\partial S_0} \right]^{-1} \left[\sum_{i=1}^{n} \frac{\partial \beta_i^T}{\partial S_0} W_i \delta\beta_i \right]$$

which is the well known, single-stage, weighted-least-squares filter without downweighting.

Discussion:

"Reasonable" values of α_j are considered to be $1 \le \alpha_j \le 3$. Referring to equation 35, e.g., α_j=1 means that the weight W_i is constant (no downweighting); α_j=2, that the weight reduces to $\frac{1}{2} W_i$ after one hour; α_j=1.03, that the weight reduces to $\frac{1}{2} W_i$ after 24 hours. "Reasonable" also implies that the selected values of α_j conform to what we know about the model and its errors. For example, since earth orbit model errors are known to be large in directions $\hat{e}_2$ and $\hat{e}_4$, the values

48. $$\alpha_2=\alpha_4=2.00, \quad \alpha_1=\alpha_3=\alpha_5=\alpha_6=1.03$$

could be considered "reasonable". Such "reasonable" choices of values for the various mission phases (earth orbit, trans-lunar, etc.) could be improved by experimenting with data from past trajectories. If necessary, these improved values could be adjusted further in real time.

The downweighting function $\alpha_j^{\,t_i-t_0}$ is used because it is simple, convenient, easily adjustable, and provides a reasonable truncation of data. It can be replaced by some better function $\gamma_j\ (t_i, t_0)$, if one is known for a particular problem.

The choice of RC was based on our knowledge of the model and its errors; perhaps the choice could be improved.

Although a six-element state vector was assumed in the formulation, the method could be applied just as well when estimating state vectors of some other order.

It is natural, navigationally, to designate the time of the last observation in the batch as anchor time. Downweighting could be applied to an estimate within the batch, however, by modifying definition 22:

$$A_{0i} \equiv \begin{bmatrix} \alpha_1^{-|t_i-t_0|} & & \emptyset \\ & \ddots & \\ \emptyset & & \alpha_6^{-|t_i-t_0|} \end{bmatrix}$$

Then observations taken later than anchor time would be downweighted in the same manner as earlier ones.

An experimental program of equation 43 is functioning as predicted. Runs were made on various batches of Apollo, earth-orbit data collected over time spans of 10 to 24 hours. "Reasonable" values of α_j, such as 48 above, did not affect the number of iterations, and estimates were improved. Undoubtedly, better sets of values for α_j could be found experimentally, but so far we have not done this.

As a final comment, the method described above should be looked at as another available mathematical tool which can be effective in appropriate circumstances.